[0141] associating each object with a different respective user based on the distance between each object and the respective users; and

[0142] generating user inputs for each user based on the detected changes in pose of the object associated to that user.

25. Computer software having computer executable instructions which, when executed by a computer cause the computer to perform the method of any of clauses 18 to 24.

26. A non-transitory, machine-readable storage medium which stores computer software according to clause 25.

1. A system for generating video game inputs, the system comprising:

an input unit operable to obtain images of a passive non-luminous object being held by a user;

an object detector operable to detect the passive non-luminous object being held by the user in the obtained images;

wherein the object detector is configured to detect an area in the image corresponding to the passive non-luminous object based on the pixels corresponding to the object and not based on a physical identifier that has been added to the object;

an object pose detector operable to detect changes in pose of the passive non-luminous object based on the obtained images of the object;

wherein the object pose detector is configured to detect the pose of the passive non-luminous object based on at least one of (i) a contour detection operation and (ii) the output of a machine learning model that has been trained to detect the poses of passive non-luminous objects in images; and

a user input generator operable to generate a user input based on detected changes in the pose of the passive non-luminous object, and to transmit an indication of the generated user input to a video game unit executing an instance of a video game, so as to update the display of a virtual object in the video game in accordance with the generated user input.

2. The system of claim **1**, wherein the object detector is configured to detect a contour in the obtained images corresponding to a periphery of the passive non-luminous object that is being held by the user; and

wherein the object pose detector is configured to detect changes in pose of the passive non-luminous object based on changes in least one of the orientation, position and area of the contour in the obtained images.

3. The system of claim **2**, wherein the user input generator is configured to generate different respective user inputs based on whether the orientation, position and area of the contour is detected as changing in the obtained images.

4. The system of claim **2**, wherein the object detector is configured to detect a plurality of contours in the obtained images, and to identify the largest contour as corresponding to the passive non-luminous object being held by the user.

5. The system of claim **2**, wherein the object detector is configured to obtain colour information indicating a pre-determined colour of the passive non-luminous object that a user is or intends to hold; and

wherein the object detector is configured to filter one or more colours not corresponding to the pre-determined colour from the obtained images, prior to performing the contour detection.

6. The system of claim **1**, wherein the input unit is configured to obtain images of at least two passive non-luminous objects being held by a user;

wherein the object detector is configured to detect respective contours corresponding to the at least two passive non-luminous objects in the obtained images, and determine a respective location within the image representation of each object;

wherein the object pose detector is configured to detect a change in pose of the at least two passive non-luminous objects based on changes in orientation of a line joining the detected locations; and

wherein the user input generator is operable to generate a directional user input so as to control a direction in which the virtual object is travelling in the video game, and to transmit the generated user input to the video game unit.

7. The system of claim **6**, wherein object detector is configured to determine a respective location within the image representations of each object by calculating an image moment for the pixels within the contours detected for each non-luminous object so as to determine a point or region that is representative of the centre of each object.

8. The system of claim **6**, wherein the object detector is configured to detect at least two contours associated with each passive non-luminous object being held by the user;

wherein the object detector is configured to generate a linear representation of each object based on the contours associated with that object; and

wherein the object pose detector is configured to detect a change in pose of the at least two non-luminous objects based on changes in orientation of a line intersecting the linear representations generated for each object.

9. The system of claim **6**, wherein the object pose detector is configured to detect whether the distance between the locations detected for each object is less than a threshold distance; and

wherein the user input generator is configured to generate a user input corresponding to a change in control mode in response to receiving an input from the object pose detector indicating that the distance between the locations detected for each object is less than the threshold distance.

10. The system of claim **1**, wherein the object detector comprises a cascade classifier trained with images of passive non-luminous objects.

11. The system of claim **1**, wherein the object pose detector comprises a machine learning model trained to perform six-dimensional pose estimation for each passive non-luminous object detected being held by the user in the obtained images.

12. The system of claim **11**, wherein the obtained images comprise at least two passive non-luminous objects being held by a user;

wherein the object pose detector is configured to detect a plurality of keypoints for each passive non-luminous object being held by the user; and

wherein the user input generator is configured to generate a user input based on a detection of the distance between at least some of the keypoints for each object being detected as greater or less than a threshold distance.